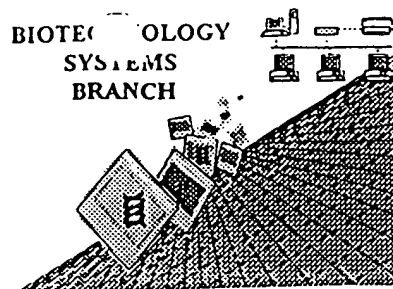# RAW SEQUENCE LISTING
ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/825,242 A

Source: OIPE

BEST AVAILABLE COPY

Date Processed by STIC: 8/23/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN-2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

OIPE

RAW SEQUENCE LISTING   DATE: 08/23/2001
PATENT APPLICATION: US/09/825,242A   TIME: 08:51:56

Input Set : A:\19496-18.app
Output Set: N:\CRF3\08162001\I825242A.raw

```
    3 <110> APPLICANT: Eisenberg, Stephen P.
    4         Case, Casey C.
    5         Cox III, George N.
    6         Jamieson, Andrew
    7         Rebar, Edward J.
    8         Sangamo Biosciences, Inc.
   10 <120> TITLE OF INVENTION: Selection of Sites for Targeting by Zinc Finger
   11         Proteins and Methods of Designing Zinc Finger Proteins
   12         to Bind to Preselected Sites
   14 <130> FILE REFERENCE: 019496-001800US
C--> 16 <140> CURRENT APPLICATION NUMBER: US/09/825,242A
C--> 17 <141> CURRENT FILING DATE: 2001-04-04
   19 <160> NUMBER OF SEQ ID NOS: 97
   21 <170> SOFTWARE: PatentIn Ver. 2.1
   23 <210> SEQ ID NO: 1
   24 <211> LENGTH: 25
   25 <212> TYPE: PRT
   26 <213> ORGANISM: Artificial Sequence
   28 <220> FEATURE:
   29 <223> OTHER INFORMATION: Description of Artificial Sequence:exemplary motif
   30         characterizing the C-2H-2 class of zinc finger
   31         proteins (ZFP)
   33 <220> FEATURE:
   34 <221> NAME/KEY: MOD_RES
   35 <222> LOCATION: (1)..(25)
   36 <223> OTHER INFORMATION: Xaa = any amino acid
   38 <220> FEATURE:
   39 <221> NAME/KEY: MOD_RES
   40 <222> LOCATION: (4)..(5)
   41 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
   43 <220> FEATURE:
   44 <221> NAME/KEY: MOD_RES
   45 <222> LOCATION: (23)..(24)
   46 <223> OTHER INFORMATION: Xaa = any amino acid, may be present or absent
   48 <400> SEQUENCE: 1
W--> 49 Cys Xaa Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
   50   1               5                  10                  15
W--> 52 Xaa Xaa His Xaa Xaa Xaa Xaa Xaa His
   53             20                  25
   56 <210> SEQ ID NO: 2
   57 <211> LENGTH: 5
   58 <212> TYPE: PRT
   59 <213> ORGANISM: Artificial Sequence
   61 <220> FEATURE:
   62 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
   64 <400> SEQUENCE: 2
   65 Thr Gly Glu Lys Pro
```

RAW SEQUENCE LISTING   DATE: 08/23/2001
PATENT APPLICATION: US/09/825,242A   TIME: 08:51:56

Input Set : A:\19496-18.app
Output Set: N:\CRF3\08162001\I825242A.raw

```
66   1              5
69 <210> SEQ ID NO: 3
70 <211> LENGTH: 5
71 <212> TYPE: PRT
72 <213> ORGANISM: Artificial Sequence
74 <220> FEATURE:
75 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
77 <400> SEQUENCE: 3
78 Gly Gly Gly Gly Ser
79   1              5
82 <210> SEQ ID NO: 4
83 <211> LENGTH: 8
84 <212> TYPE: PRT
85 <213> ORGANISM: Artificial Sequence
87 <220> FEATURE:
88 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
90 <400> SEQUENCE: 4
91 Gly Gly Arg Arg Gly Gly Gly Ser
92   1              5
95 <210> SEQ ID NO: 5
96 <211> LENGTH: 9
97 <212> TYPE: PRT
98 <213> ORGANISM: Artificial Sequence
100 <220> FEATURE:
101 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
103 <400> SEQUENCE: 5
104 Leu Arg Gln Arg Asp Gly Glu Arg Pro
105   1              5
108 <210> SEQ ID NO: 6
109 <211> LENGTH: 12
110 <212> TYPE: PRT
111 <213> ORGANISM: Artificial Sequence
113 <220> FEATURE:
114 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
116 <400> SEQUENCE: 6
117 Leu Arg Gln Lys Asp Gly Gly Gly Ser Glu Arg Pro
118   1              5                 10
121 <210> SEQ ID NO: 7
122 <211> LENGTH: 16
123 <212> TYPE: PRT
124 <213> ORGANISM: Artificial Sequence
126 <220> FEATURE:
127 <223> OTHER INFORMATION: Description of Artificial Sequence:peptide linker
129 <400> SEQUENCE: 7
130 Leu Arg Gln Lys Asp Gly Gly Gly Ser Gly Gly Gly Ser Glu Arg Pro
131   1              5                 10            15
134 <210> SEQ ID NO: 8
135 <211> LENGTH: 85
136 <212> TYPE: PRT
```

137 <213> ORGANISM: Mus sp.
139 <220> FEATURE:
140 <223> OTHER INFORMATION: DNA binding domain of mouse transcription factor
141   Zif268
143 <400> SEQUENCE: 8
144 Tyr Ala Cys Pro Val Glu Ser Cys Asp Arg Arg Phe Ser Arg Ser Asp
145 1   5   10   15
147 Glu Leu Thr Arg His Ile Arg Ile His Thr Gly Gln Lys Pro Phe Gln
148   20   25   30
150 Cys Arg Ile Cys Met Arg Asn Phe Ser Arg Ser Asp His Leu Thr Thr
151   35   40   45
153 His Ile Arg Thr His Thr Gly Glu Lys Pro Phe Ala Cys Asp Ile Cys
154   50   55   60
156 Gly Arg Lys Phe Ala Arg Ser Asp Glu Arg Lys Arg His Thr Lys Ile
157 65   70   75   80
159 His Leu Arg Gln Lys
160   85
163 <210> SEQ ID NO: 9
164 <211> LENGTH: 94
165 <212> TYPE: PRT
166 <213> ORGANISM: Artificial Sequence
168 <220> FEATURE:
169 <223> OTHER INFORMATION: Description of Artificial Sequence:amino acids
170   531-624 in Sp-1 transcription factor
172 <400> SEQUENCE: 9
173 Pro Gly Lys Lys Lys Gln His Ile Cys His Ile Gln Gly Cys Gly Lys
174 1   5   10   15
176 Val Tyr Gly Lys Thr Ser His Leu Arg Ala His Leu Arg Trp His Thr
177   20   25   30
179 Gly Glu Arg Pro Phe Met Cys Thr Trp Ser Tyr Cys Gly Lys Arg Phe
180   35   40   45
182 Thr Arg Ser Asp Glu Leu Gln Arg His Lys Arg Thr His Thr Gly Glu
183   50   55   60
185 Lys Lys Phe Ala Cys Pro Glu Cys Pro Lys Arg Phe Met Arg Ser Asp
186 65   70   75   80
188 His Leu Ser Lys His Ile Lys Thr His Gln Asn Lys Lys Gly
189   85   90
192 <210> SEQ ID NO: 10
193 <211> LENGTH: 98
194 <212> TYPE: PRT
195 <213> ORGANISM: Artificial Sequence
197 <220> FEATURE:
198 <223> OTHER INFORMATION: Description of Artificial Sequence:Sp-1
199   transcription factor consensus sequence
201 <400> SEQUENCE: 10
202 Met Glu Lys Leu Arg Asn Gly Ser Gly Asp Pro Gly Lys Lys Lys Gln
203 1   5   10   15
205 His Ala Cys Pro Glu Cys Gly Lys Ser Phe Ser Lys Ser Ser His Leu
206   20   25   30

RAW SEQUENCE LISTING   DATE: 08/23/2001
PATENT APPLICATION: US/09/825,242A   TIME: 08:51:56

Input Set : A:\19496-18.app
Output Set: N:\CRF3\08162001\I825242A.raw

```
208 Arg Ala His Gln Arg Thr His Thr Gly Glu Arg Pro Tyr Lys Cys Pro
209         35                40                45
211 Glu Cys Gly Lys Ser Phe Ser Arg Ser Asp Glu Leu Gln Arg His Gln
212         50              55                60
214 Arg Thr His Thr Gly Glu Lys Pro Tyr Lys Cys Pro Glu Cys Gly Lys
215 65                70                75                80
217 Ser Phe Ser Arg Ser Asp His Leu Ser Lys His Gln Arg Thr His Gln
218                 85              90              95
220 Asn Lys
223 <210> SEQ ID NO: 11
224 <211> LENGTH: 10
225 <212> TYPE: DNA
226 <213> ORGANISM: Artificial Sequence
228 <220> FEATURE:
229 <223> OTHER INFORMATION: Description of Artificial Sequence:natural Zif268
230         binding site
232 <400> SEQUENCE: 11
233 gcgtgggcgc                                                      10
236 <210> SEQ ID NO: 12
237 <211> LENGTH: 10
238 <212> TYPE: DNA
239 <213> ORGANISM: Artificial Sequence
241 <220> FEATURE:
242 <223> OTHER INFORMATION: Description of Artificial Sequence:target site
243         containing three D-able subsites
245 <400> SEQUENCE: 12
```

W--> 246 gcntgnggnn ───────────────────────

see item #9 on ERROR SUMMARY sheet   10

```
249 <210> SEQ ID NO: 13
250 <211> LENGTH: 10
251 <212> TYPE: DNA
252 <213> ORGANISM: Artificial Sequence
254 <220> FEATURE:
255 <223> OTHER INFORMATION: Description of Artificial Sequence:target site
256         with two overlapping D-able subsites
258 <400> SEQUENCE: 13
```

W--> 259 nngkngknnn   10

```
262 <210> SEQ ID NO: 14
263 <211> LENGTH: 10
264 <212> TYPE: DNA
265 <213> ORGANISM: Artificial Sequence
267 <220> FEATURE:
268 <223> OTHER INFORMATION: Description of Artificial Sequence:target site
269         with three overlapping D-able subsites
271 <400> SEQUENCE: 14
```

W--> 272 nngkngkngk   10

```
275 <210> SEQ ID NO: 15
276 <211> LENGTH: 22
277 <212> TYPE: DNA
278 <213> ORGANISM: Artificial Sequence
```

RAW SEQUENCE LISTING   DATE: 08/23/2001
PATENT APPLICATION: US/09/825,242A   TIME: 08:51:56

Input Set : A:\19496-18.app
Output Set: N:\CRF3\08162001\I825242A.raw

```
280 <220> FEATURE:
281 <223> OTHER INFORMATION: Description of Artificial Sequence:target site DNA
282         motif searched by protocol 1
284 <220> FEATURE:
285 <221> NAME/KEY: modified_base
286 <222> LOCATION: (1)..(22)
287 <223> OTHER INFORMATION: n = g, a, c or t
289 <220> FEATURE:
290 <221> NAME/KEY: modified_base
291 <222> LOCATION: (10)..(12)
292 <223> OTHER INFORMATION: n = g, a, c or t, may be present or absent
294 <400> SEQUENCE: 15
```
W--> 295 gnggnngnnn nngnggnngn nn   22
```
298 <210> SEQ ID NO: 16
299 <211> LENGTH: 23
300 <212> TYPE: DNA
301 <213> ORGANISM: Artificial Sequence
303 <220> FEATURE:
304 <223> OTHER INFORMATION: Description of Artificial Sequence:target site DNA
305         motif searched by protocol 1
307 <220> FEATURE:
308 <221> NAME/KEY: modified_base
309 <222> LOCATION: (1)..(23)
310 <223> OTHER INFORMATION: n = g, a, c or t
312 <220> FEATURE:
313 <221> NAME/KEY: modified_base
314 <222> LOCATION: (11)..(13)
315 <223> OTHER INFORMATION: n = g, a, c or t, may be present or absent
317 <400> SEQUENCE: 16
```
W--> 318 gnggnngnnn nnngnggnng nnn   23
```
321 <210> SEQ ID NO: 17
322 <211> LENGTH: 22
323 <212> TYPE: DNA
324 <213> ORGANISM: Artificial Sequence
326 <220> FEATURE:
327 <223> OTHER INFORMATION: Description of Artificial Sequence:target site DNA
328         motif searched by protocol 1
330 <220> FEATURE:
331 <221> NAME/KEY: modified_base
332 <222> LOCATION: (1)..(22)
333 <223> OTHER INFORMATION: n = g, a, c or t
335 <220> FEATURE:
336 <221> NAME/KEY: modified_base
337 <222> LOCATION: (10)..(12)
338 <223> OTHER INFORMATION: n = g, a, c or t, may be present or absent
340 <400> SEQUENCE: 17
```
W--> 341 gnggnngnnn nngnngnggn nn   22
```
344 <210> SEQ ID NO: 18
345 <211> LENGTH: 23
```

FYI:

Use of 'n' and/or 'Xaa' has been detected in the Sequence Listing. Review the
Sequence Listing to insure a corresponding explanation is presented in the
<220> to <223> fields of each sequence using 'n' or 'Xaa'.

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/825,242A

DATE: 08/23/2001
TIME: 08:51:57

Input Set : A:\19496-18.app
Output Set: N:\CRF3\08162001\I825242A.raw

```
L:16 M:270 C: Current Application Number differs, Replaced Current Application Number
L:17 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:49 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:52 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:246 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:12
L:246 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:12
L:246 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:12
L:259 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:13
L:259 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:13
L:259 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13
L:272 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:14
L:272 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:14
L:272 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14
L:295 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:318 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16
L:341 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:17
L:364 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18
L:387 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:19
L:410 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:20
L:433 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21
L:456 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:479 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23
L:502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24
L:525 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:25
L:548 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:26
L:571 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:27
L:594 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:28
L:617 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:29
L:640 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30
L:663 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31
L:686 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32
L:704 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:33
L:722 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:34
L:740 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:35
L:763 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:36
L:786 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:37
L:809 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:38
L:832 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:39
L:855 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:40
L:878 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:41
L:901 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:42
L:924 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:43
L:947 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:44
L:970 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:45
L:993 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:1016 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47
L:1039 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:48
L:1062 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:49
```

```
L:1085 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:50
L:1108 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:51
L:1131 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:52
L:1154 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:53
L:1172 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:54
L:1190 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:55
L:1208 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:56
L:1231 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:57
L:1254 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:58
L:1277 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:59
```